DOCUMENT RESUME

ED 118 453 SE 020 309

AUTHOR Nash, A. H.
TITLE Improving Accuracy of Assessment Procedures.
PUB DATE Jun 75
NOTE 28p.; Paper presented at the Annual Meeting of the American Society for Engineering Education (Colorado State University, Ft. Collins, Colorado, June 16-19, 1975); Occasional marginal legibility in Figure 3-7

EDRS PRICE MF-$0.83 HC-$2.06 Plus Postage
DESCRIPTORS Achievement; College Science; *Computer Programs; Engineering Education; *Evaluation; *Grades (Scholastic); Higher Education; Science Education; *Student Evaluation; *Vocational Schools
IDENTIFIERS *Western Australian Institute of Technology

ABSTRACT

A review of the grading practices of various departments in the Western Australian Institute of Technology is the topic of this paper. The study was initiated in 1969, when an examination of scores given by various departments revealed a large year-to-year fluctuation. It was noted that some departments consistently graded higher than others. A historical account of the study, with graphs showing the grading inequities, is provided. The inception of a "standard grading distribution" and problems which resulted from its implementation are also described. Due to the number of arbitration situations in which students questioned their grades, a computer program was developed which made examination and student academic information available to arbitrators. The author states that his addition of an easily accessible information network has greatly increased the efficiency of the grading system. (CP)

EVENT NO. 2530

AMERICAN SOCIETY FOR ENGINEERING EDUCATION

ANNUAL CONFERENCE, JUNE 16 - 19, 1975.

COLORADO STATE UNIVERSITY

FT. COLLINS, CO. 80521

# IMPROVING ACCURACY OF ASSESSMENT PROCEDURES

DR. A.H. NASH

DEAN, SCHOOL OF ENGINEERING AND SURVEYING,

WESTERN AUSTRALIAN INSTITUTE OF TECHNOLOGY.

# IMPROVING ACCURACY OF ASSESSMENT PROCEDURES

## The Use of Students' Academic Histories and the Computer to Improve the Accuracy of Assessing Academic Performance of Students.

---


SUMMARY

Several years ago investigations at the Western Australian Institute of Technology showed that scores in certain subjects with large student numbers varied widely from year to year. It was further noted that, in any one year, scores obtained in certain subjects varied widely from those in allied subjects where similar performance could be expected.

This caused concern. Were the variations due to faulty teaching, faulty assessment techniques, a faulty curriculum, or even faulty students.

As a result the Institute reviewed its scoring processes and brought down guidelines for score distributions. These were intended to identify problem subjects so that the causes of the troubles could be determined and appropriate remedial action taken.

Some lecturers accepted the guidelines as they were intended and

assessment processes improved. Others considered the guidelines an infringement of "academic freedom" and reacted accordingly. Yet again a third group adjusted scores to conform to the distribution without giving real thought to the educational reasons for doing this.

One of the problems in utilising the considerable volume of available material to improve assessment was a time limitation between examinations and the issuing of results. Subsequently a computer program was prepared so that the information required could be provided in the available time span.

This has had considerable impact. Many of the opponents of the guideline scheme have now changed their attitude and a marked improvement of the assessment procedure has resulted.

Details of the development of the system, the computer printouts made available, and the benefits observed are considered in the paper.

## 1 INTRODUCTION

The author's experience with student assessment at the tertiary level in the USA has been restricted to the post-graduate level. A point that impressed was the expediency with which results were processed (a post card indicating the grade achieved usually arrived within 48 hours) and the fact that grades were generally higher than

those given in Australia.

At the undergraduate level in Australia, and in particular at the Western Australian Institute of Technology, a detailed analysis of students' performance is undertaken. Before any results are issued the performance of students in all subjects is recorded and a Board of Examiners reviews overall performance. If the overall performance of a student is good but that in a particular subject is below the pass level he may be granted a conceded pass or at least a supplementary examination in that subject.

The advantage of the system is that it looks at the overall performance of a student in his field. Unfortunately the procedure also has short-comings. In a 1969[1] study it was observed that the year by year score variation in some subjects with large student populations was much greater than expected. These changes could not be accepted, particularly when the score variations in allied subjects with similar students and numbers were comparatively small. Figure 1 illustrates the score range obtained in one such subject and in another subject.

---

[1] Ref. 1, p. 38

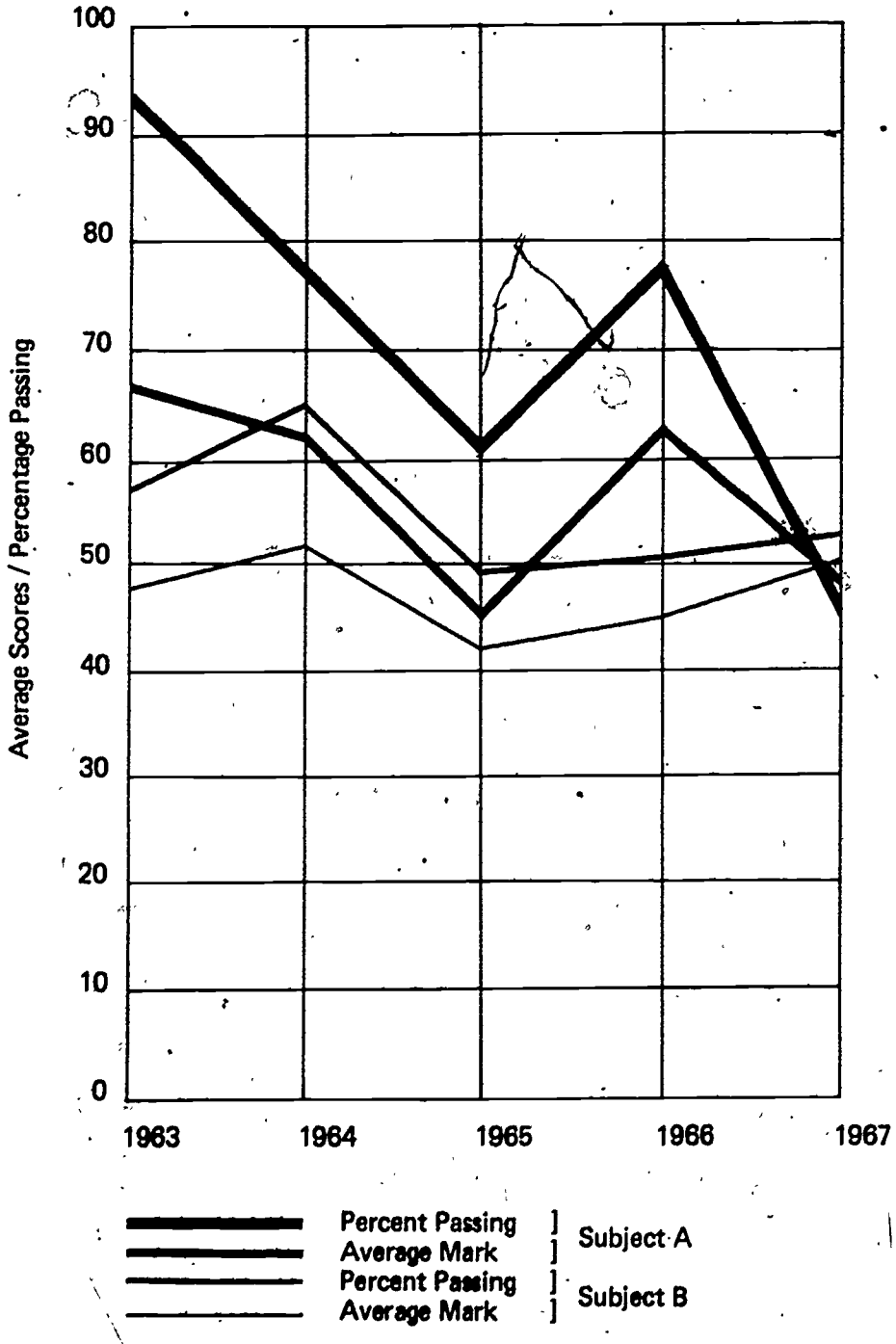


**FIGURE 1—VARIATION OF AVERAGE SCORES AND PERCENTAGE OF STUDENTS PASSING IN PARTICULAR ANNUAL EXAMINATIONS**

What was the cause? Faulty curriculum, faulty assessment procedure, faulty teaching, student malperformance or something else?

Institute senior management decided that steps should be taken to eliminate or at least minimise the problem.

## 2 THE PROBLEMS

It is reasonable to expect that in large classes, the range of ability should be reasonably constant from year to year. Thus assessment scores should also be reasonably constant if examinations measure the students' ability accurately[2] and teaching is consistent from year to year.

Likewise it was expected that examination score distribution Institute wide would be reasonably constant from year to year. This could provide a distribution pattern for individual subjects.[3]

The Institute distribution of scores was determined for 1967 and is as in figure 2.

---

[2] With small classes the range of student ability could vary considerably from year to year giving rise to varying examination scores.

[3] The student population of the Institute is currently of the order of 10,000.

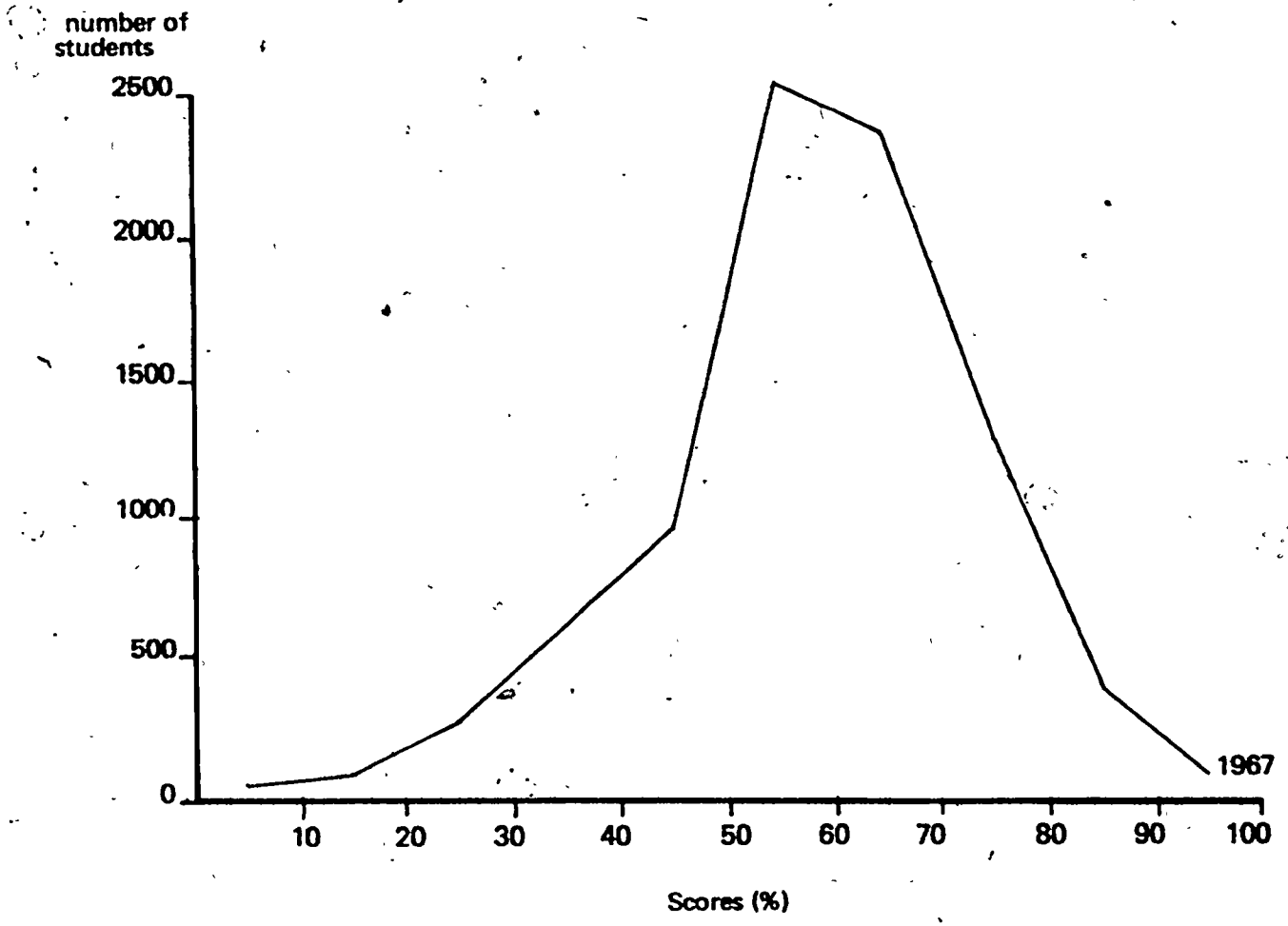


FIGURE 2—ACTUAL DISTRIBUTION OF SCORES INSTITUTE WIDE 1967

After wide-ranging discussions the Institute's management group accepted the following frequency distribution guidelines for examiners. These were based on the information in figure 2 and other material available.

TABLE 1

GUIDELINE DISTRIBUTION OF SCORES IN INSTITUTE EXAMINATIONS

| Examination Score Range | Grade | Percentage Population in Range |
|---|---|---|
| 75 - 100 | A | 0 - 25 |
| 65 - 74 | B | 10 - 40 |
| 50 - 64 | C | 30 - 65 |
| 40 - 49 | F[4] | 0 - 20 |
| 20 - 39 | G | 0 - 10 |
| 0 - 19 | R | excluded |

These figures reflect a desire to improve the performance of that section of the "tail" which performed reasonably well but to exclude the group who scored in the 0 - 19 range as it was considered these were not really sincere students. Thus the percentages in column 3 of Table 1 were to be applied to the sample remaining after the "R" group had been excluded.

---

[4] Students who scored less than 50 failed the examination.

## 3 IMPLEMENTATION

The proposals were implemented in 1970. Where examiners had reasonable background in educational theory, thus appreciating the limitations associated with assessment procedures and having knowledge of scaling techniques, there was widespread acceptance of the scheme.

However, a second group asserted their right to "protect the profession" and fail students when their raw score was below 50 per cent. This could result in a student who scored 49% being failed even though the highest score achieved in the examination may have been 60% and where scores of the same students in similar subjects ranged to maximums in the nineties.

A third group took the easy way out by adjusting scores to conform to the guidelines without considering why they were doing so. They were simply conforming with management policy.

Thus the first group of lecturers and their students benefited from the program but the same could not be said of the others.

The second group in particular caused considerable concern and it was decided that if the examiner and head of department could not agree with the Dean regarding the need to adjust particular scores the matter would be referred to an arbitration group.

The proceedings of the arbitration group indicated that, to some examiners, the scoring process was precise and objective. It was of concern to realise that the inherent limitations and problems were not appreciated.

A typical case will illustrate the procedure at an arbitration hearing. In a particular subject an unduly large percentage of students had failed. The examiner, supported by his head of department, argues that he is justified in retaining this pass rate because the students had failed in other subjects. The following facts are brought forward by the Dean to justify a revision of scores.

i. The particular student group had similar course entry scores to engineering students in other departments. These other students had performed well.

ii. The performance of the student group in common subjects such as mathematics and physics was similar to that of other engineering students.

iii. The scores for the group under consideration were depressed in comparison with those in previous years.

iv. Investigation of performance on individual questions in the examination showed there had been two questions with universally poor scores.

The department had not obtained this information and hence it could not make a true evaluation. In fact the student definitely was not at fault!

Unfortunately an appeal court is not the place for major adjustments to be made. Prejudices prevail and it is difficult to effect changes. Consequently though some improvements were made, much was left to be desired.

4. UPGRADING THE SYSTEM

The policy as detailed continued through 1971 and 1972 with fairly regular confrontation between Deans, Heads of Departments and Examiners. It became apparent that much more could be achieved if the Boards of Examiners could have before them the detailed analysis of performance and historical records that were laboriously produced manually for the confrontation sessions.

The computer was the logical machine to expedite the provision of this information. It was already used to produce a printout showing the performance of each student in all subjects sat in an examination.

Mr. S. Nowak, a member of the Mechanical Engineering Department, was commissioned to investigate the matter and prepare appropriate computer programs.

Experience had shown that the following deficiencies were associated with the existing assessing scheme:

i. particular examiners considered student scores in their subject in isolation.

ii. The performance of individual students as well as that of the group in a particular subject were not compared with performances in related subjects.

iii. Generally a student's performance in a particular subject was not compared with the overall performance of students in his class (or his group where there was more than one teaching group).

iv. No analysis was conducted to ascertain the relative performance of all students on individual questions in an examination paper.

v. No attempt was made to determine if there had been consistency of syllabus treatment where more than one lecturer handled the subject.

vi. The earlier performance of a student and his course entrance scores were not readily available to Boards of Examiners when assessing the student's performance.

## 5 THE UPGRADED SYSTEM

The computerised system analysing students' performance was subdivided into three inter-related parts:

i. The preparation by the computer of mark sheets.

ii. The analysis of examination results from the information submitted on mark sheets.

iii. Printouts providing information on which to judge the students present and future status at the Institute.

The three separate computer programs prepared for these steps are now considered in detail.

### 5.1 Mark Sheets

The computer prints out separate sheets for each subject being examined. These list the enrolled students alphabetically and also provide columns for scores obtained in each question in the end of semester examination, semester scores for laboratory work, assignments and tests, and information as to whether the student is attempting the subject for the first, second or third time. The sheets are custom made to fit the requirements of particular departments, a typical sample partially completed, is shown in figure 3.

74/11/12 WESTERN AUSTRALIAN INSTITUTE OF TECHNOLOGY

SCHOOL OF ENGINEERING & SURVEYING SECOND SEMESTER EXAMINATIONS

SUBJECT 7187 PAPER 1 EXAMINER.....

| SUBJ | P | STUDENT | DEP | CRSE | GRP | Q1 | Q2 | Q3 | Q4 | Q5 | Q6 | Q7 | Q8 | Q9 | Q10 | Q11 | Q12 | ASG | LAB | COM |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 7187 | 1 | 742375E | 11 | 712 | 2 | 4 | 25 | 5 | | 12 | | | | | | | | 95 | | 10 |
| 7187 | 1 | 7436296 | 11 | 712 | 2 | 2 | 11 | 7 | | 20 | | | | | | | | 75 | | 10 |
| 7187 | 1 | 7427898 | 11 | 712 | 2 | 0.1 | 5 | | 22 | 13 | | | | | | | | 80 | | 10 |
| 7187 | 1 | 742245A | 11 | 712 | | | | | | | | | | | | | | | | |
| 7187 | 1 | 743991A | 11 | 712 | | | | | | | | | | | | | | | | |
| 7187 | 1 | 7[illegible]192B | 11 | 712 | | | | | | | | | | | | | | | | |
| 7187 | 1 | 743442F | 11 | 712 | | | | | | | | | | | | | | | | |
| 7187 | 1 | 742927H | 11 | 712 | | | | | | | | | | | | | | | | |
| 7187 | 1 | 731971E | 11 | 712 | | | | | | | | | | | | | | | | |
| 7187 | 1 | 741041J | 11 | 712 | | | | | | | | | | | | | | | | |
| 7187 | 1 | 743618C | 11 | 712 | | | | | | | | | | | | | | | | |
| 7187 | 1 | 742413C | 11 | 712 | | | | | | | | | | | | | | | | |

NAMES OF STUDENTS HAVE BEEN DELETED TO PRESERVE ANONYMITY.

15

13

FIG 3

The column GRP is used to indicate the group number. Thus if the class is divided into four teaching groups these would be identified at GRP 1, 2, 3 and 4.

Question scores are entered in columns Q1, Q2 and so on, whilst the ASG and LAB columns take assignment and laboratory scores. The final column, COM is used to record the number of attempts a student has had in a particular subject.

It has been found useful to know how many students actually attempted a particular question as this had obvious implications regarding its difficulty or the effectiveness of the treatment of that particular part of the syllabus. Consequently examiners were instructed to enter a mark of 0.1 where a student had attempted a question but obtained zero marks, and leave a blank if the question was not attempted.

Provided teaching departments submit the information in time mark sheets can be produced at the beginning of a semester. They are then particularly valuable where continuous assessment is being used. Alternately the sheets can be provided at a later date as required. Irrespective of when the sheets are provided the relevant material must be entered and returned to the computer center for processing immediately after the final semester examinations.

## 5.2 Analysis of Examination Results

Prior to the examination the computer has been advised of the number of questions in the particular examination paper, the number of questions to be attempted, the maximum marks for each question, and for assignments, laboratory work and continuous assessment programs. The weighting to be given to the end of semester examination and the other components of the assessment procedure are also programmed in. Thus when the computer receives the information from subject mark sheets, via punched cards, it can immediately print out the raw score and corresponding grade (0 - 9) for each student as in figure 4. Additional information is also provided then as shown in figure 5, which details the following.

i. The average score and the standard deviation for the subject.

ii. The distribution of grades for students new to the subject and for those repeating for the first and second time.

iii. The performance of the class as a whole on various individual questions.

An additional similar printout gives the information of items i and ii above for the various teaching groups within the class.

FIG 4

SEMESTER ASSESSMENT 29/11/7***

PAPER 1
GROUP (2)

| | STUDENT NAME | 01 | 02 | 03 | 04 | 05 | 06 | 07 | EVAL | ASSG | LAB | RESULT | GRADE | CU |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | | 4.0 | 25.0 | 5.0 | | 12.0 | | | 50.2 | 27.5 | | 61.7% | (6) | 10 |
| 2 | | 12.0 | 11.0 | 7.0 | | 20.0 | | | 45.0 | 22.5 | | 50.5% | (5) | 10 |
| 3 | | | 0.1 | 5.0 | 22.0 | 13.0 | | | 43.1 | 24.0 | | 52.1% | (5) | 10 |
| 4 | | 19.0 | 18.0 | 8.0 | | 12.0 | | | 48.1 | 19.0 | | 62.1% | (6) | 10 |
| 5 | | 6.0 | 2.0 | | 1.0 | 13.0 | | | 15.4 | 23.5 | | 40.9% | (4) | 10 |
| 6 | | | 8.0 | | 12.0 | 12.0 | | | 22.4 | 21.0 | | 43.4% | (4) | 11 |
| 7 | | 8.0 | 9.0 | 8.0 | 8.0 | | | | 23.4 | 24.0 | | 47.1% | (4) | 11 |
| 8 | | | 21.0 | 5.0 | 11.0 | 21.0 | | | 41.5 | 27.5 | | 69.1% | (6) | 10 |
| 9 | | 7.0 | 7.0 | 5.0 | 24.0 | | | | 36.1 | 22.5 | | 52.6% | (5) | 10 |
| 10 | | 6.0 | | 0.1 | 22.0 | 19.0 | | | 33.4 | 19.0 | | 51.0% | (5) | 11 |
| 11 | | | 22.0 | 5.0 | 22.0 | 17.0 | | | 44.2 | 13.5 | | 58.7% | (5) | 10 |
| 12 | | 7.0 | 10.0 | 11.0 | | 20.0 | | | 38.6 | 27.0 | | 60.6% | (6) | 10 |
| 13 | | | 1.0 | 5.0 | 22.0 | 11.0 | | | 27.5 | 19.0 | | 45.5% | (4) | 11 |
| 14 | | | 25.0 | 8.0 | 25.0 | 13.0 | | | 49.7 | 24.0 | | 73.7% | (7) | 10 |
| 15 | | | 6.0 | 8.0 | 25.0 | 18.0 | | | 39.0 | 27.5 | | 65.4% | (6) | 10 |
| 16 | | | 21.0 | 10.0 | 25.0 | 24.0 | | | 62.3 | 27.0 | | 89.3% | (6) | 10 |
| 17 | | | 8.0 | 7.0 | 8.0 | 10.0 | | | 23.1 | 27.0 | | 50.1% | (5) | 10 |
| 18 | | 12.0 | 10.0 | 8.0 | 11.0 | | | | 35.0 | 24.0 | | 59.0% | (5) | 11 |
| 19 | | 12.0 | 0.1 | 23.0 | 11.0 | | | | 32.3 | 21.0 | | 53.5% | (5) | 11 |
| 20 | | 4.0 | 0.1 | | 15.0 | 16.0 | | | 24.6 | 21.0 | | 45.6% | (4) | 11 |
| 21 | | | 21.0 | 5.0 | 0.1 | 22.0 | | | 33.7 | 25.5 | | 59.2% | (5) | 10 |
| 22 | | 7.0 | 24.0 | 25.0 | | 9.0 | | | 44.3 | 21.0 | | 65.8% | (6) | 10 |
| 23 | | 14.0 | | 0.1 | | 7.0 | | | 7.8 | 24.0 | | 31.8% | (5) | 10 |
| 24 | | | 11.0 | 7.0 | 11.0 | 20.0 | | | 34.5 | 21.0 | | 55.3% | (5) | 10 |
| 25 | | | 4.0 | 0.0 | 22.0 | 10.0 | | | 31.5 | 25.5 | | 57.0% | (5) | 10 |
| 26 | | 17.0 | 25.0 | 25.0 | | 23.0 | | | 63.0 | 27.0 | | 90.0% | (9) | 10 |

NAMES OF STUDENTS HAVE BEEN DELETED TO PRESERVE ANONYMITY.

***SUMMARY SELECTED ASSESSMENT 29/11/76***

PAGE 1

DISTRIBUTION OF GRADES NEW ENTRY STUDENTS

| | | |
|---|---|---|
| GRADE 1 = 1.0 ( 1.6%) | GRADE 2 = 2.0 ( 3.2%) | GRADE 3 = 2.0 ( 3.2%) |
| GRADE 4 = 5.0 ( 7.9%) | GRADE 5 =15.0 (23.8%) | GRADE 6 =14.0 (22.2%) |
| GRADE 7 = 8.0 (12.7%) | GRADE 8 = 2.0 ( 3.2%) | GRADE 9 = 2.0 ( 3.2%) |

DISTRIBUTION OF GRADES REPEAT STUDENTS(FIRST TIME)

| | | |
|---|---|---|
| GRADE 1 = 0.0 ( 0.0%) | GRADE 2 = 0.0 ( 0.0%) | GRADE 3 = 0.0 ( 0.0%) |
| GRADE 4 = 0.0 ( 0.0%) | GRADE 5 = 0.0 ( 0.0%) | GRADE 6 = 0.0 ( 0.0%) |
| GRADE 7 = 0.0 ( 0.0%) | GRADE 8 = 0.0 ( 0.0%) | GRADE 9 = 0.0 ( 0.0%) |

DISTRIBUTION OF GRADES REPEAT STUDENTS(SECOND TIME)

| | | |
|---|---|---|
| GRADE 1 = 0.0 ( 0.0%) | GRADE 2 = 1.0 ( 1.6%) | GRADE 3 = 0.0 ( 0.0%) |
| GRADE 4 = 4.0 ( 6.3%) | GRADE 5 = 3.0 ( 4.8%) | GRADE 6 = 4.0 ( 6.3%) |
| GRADE 7 = 0.0 ( 0.0%) | GRADE 8 = 0.0 ( 0.0%) | GRADE 9 = 0.0 ( 0.0%) |

TOTAL DISTRIBUTION OF GRADES

| | | |
|---|---|---|
| GRADE 1 = 1.0 ( 1.6%) | GRADE 2 = 3.0 ( 4.8%) | GRADE 3 = 2.0 ( 3.2%) |
| GRADE 4 = 9.0 (14.3%) | GRADE 5 =18.0 (28.6%) | GRADE 6 =18.0 (28.6%) |
| GRADE 7 = 8.0 (12.7%) | GRADE 8 = 2.0 ( 3.2%) | GRADE 9 = 2.0 ( 3.2%) |

DISTRIBUTION OF MARKS FOR EACH QUESTION

| MARKS SCORED | Q1 | Q2 | Q3 | Q4 | Q5 |
|---|---|---|---|---|---|
| 0.0 | 40.0(63.5%) | 7.0(11.1%) | 9.0(14.3%) | 9.0(14.3%) | 5.0( 7.9%) |
| 0.1-- 6.0 | 7.0(11.1%) | 13.0(20.6%) | 15.0(23.8%) | 6.0( 9.5%) | 3.0( 4.8%) |
| 6.0--12.0 | 10.0(15.9%) | 16.0(25.4%) | 29.0(46.0%) | 10.0(15.9%) | 12.0(19.0%) |
| 12.0--18.0 | 4.0( 6.3%) | 7.0(11.1%) | 2.0( 3.2%) | 17.0(27.0%) | 16.0(25.4%) |
| 18.0--25.0 | 2.0( 3.2%) | 20.0(31.7%) | 8.0(12.7%) | 21.0(33.3%) | 27.0(42.9%) |
| TOTAL | 63.0 | 63.0 | 63.0 | 63.0 | 63.0 |

AV.RESULT FOR SUBJECT =57.912 ST.DEV. FOR THE SUBJECT =15.873

FIG 5

Figure 6 shows an alphabetical list of all students in the subject together with their final raw score and corresponding grade. Provision is made for amendments or recommendation by the Board of Examiners. The final print out prepared from the amended figure 6 gives an alphabetical list containing the final approved grade.

Some comment regarding the use of this information is appropriate.

Figure 5 indicates that question one was unpopular as it was avoided by an unexpectedly high portion of the population. It also shows that performance on this question was generally poor. The examiners are thus alerted to investigate it. Is this topic traditionally unpopular - and if so what can be done about it? Or were the results due to a poorly set or ambiguous question? Yet again were they due to poor preparation of all students in the subject or of those in one or more groups within the class?

Having identified the cause appropriate remedial action can be taken for the present and the future. For example if the difficulty is due to a poorly set examination question some scaling of scores is indicated. If the difficulty is related to the teaching of the group as a whole some form of scaling along with a review of the relevant curriculum section and teaching method should be undertaken. Where only some groups have experienced difficulty, there is a case for differential scaling together with a review of the teaching

***FINAL RESULTS 20/11/74***

| | ****NAME**** | PAPER (1) | PAPER (2) | PAPER (3) | FIN. RES. | RECOMMENDATION |
|---|---|---|---|---|---|---|
| 1 | 731455A06663 | | | | 38.6% (3) | ---------- |
| 2 | 710101A06662 | | | | 46.7% (6) | ---------- |
| 3 | 742942A06663 | | | | 90.0% (9) | ---------- |
| 4 | 730345K06663 | | | | 57.6% (5) | ---------- |
| 5 | 721203F06661 | | | | 60.6% (6) | ---------- |
| 6 | 730323F06663 | | | | 55.3% (5) | ---------- |
| 7 | 730534K06661 | | | | 31.8% (3) | ---------- |
| 8 | 711320F06662 | | | | 54.1% (5) | ---------- |
| 9 | 731509F06663 | | | | 55.7% (5) | ---------- |
| 10 | 731276E06663 | | | | 72.1% (7) | ---------- |
| 11 | 730305G06663 | | | | 65.6% (6) | ---------- |
| 12 | 732901B06561 | | | | 59.2% (5) | ---------- |
| 13 | 700435K06661 | | | | 22.4% (2) | ---------- |
| 14 | 730613B06663 | | | | 7.1% (1) | ---------- |
| 15 | 720778A06663 | | | | 45.0% (4) | ---------- |
| 16 | 733307D06663 | | | | 47.1% (4) | ---------- |
| 17 | 722005B06661 | | | | 50.1% (5) | ---------- |
| 18 | 721503A06663 | | | | 55.3% (5) | ---------- |
| 19 | 722242C06663 | | | | 43.9% (4) | ---------- |
| 20 | 730493A06661 | | | | 50.1% (5) | ---------- |
| 21 | 730335E06662 | | | | 77.0% (7) | ---------- |
| 22 | 710181F06662 | | | | 67.0% (6) | ---------- |
| 23 | 730411A06662 | | | | 55.2% (5) | ---------- |
| 24 | 730727A06663 | | | | 43.2% (4) | ---------- |
| 25 | 721934A06661 | | | | 73.7% (7) | ---------- |
| 26 | 730917G06661 | | | | 65.4% (6) | ---------- |

NAMES OF STUDENTS HAVE BEEN DELETED TO PRESERVE ANONYMITY.

FIG 6

methods of lecturers involved. Thus it is possible to compensate for short term problems in examination scoring which would otherwise penalise the student unfairly. At the same time remedial action can be taken to avoid or minimise the problem in the future.

Questions 2 and 3 were badly answered. The examiners must determine if the questions were ambiguous or if students lacked preparation. Or can some other reason for the poor results be identified?

The distribution of grades in Figure 5 shows that 57 percent of students scores are between 50 and 70 percent and the average score of 58% for the subject appears to be reasonable. The scatter at 15.8 percent may be a little high. Why is the scatter high? Referring to figure 4 it is noted there appears to be a difference between the results obtained in the examination and the overall result. The weighted marks for assignments (maximum score of 30) indicate all students performed well on assignments but poorly in examinations (maximum score of 70). Was the weighting for examinations too high - or the marking of assignments too generous? If the assignment is a reasonable reflector of the students ability, something is obviously wrong with the examination.

In addition to the advantages discussed the system aids in achieving uniformity of standards where there are several different groups in the subject and also enables a feedback report to be prepared showing the performance of a student in relation to that of others in the subject. Finally the system permits of the validation of the test instrument used, and where it is a reusable test, to suggest areas where changes are required.

Determination of Final Grades and Student Status

In doing this, the following information is provided to examiners and boards of examiners.

i. The performance of the student as an individual (figure 4).

ii. The performance of the student within a group or class in a particular subject (figures 4 and 5).

iii. The performance of the student in all subjects in the semester (figure 7).

Additionally, earlier history records of the student are printed out (figure 7).

To indicate the relative standing of various students in the subject by comparison with the year reference mark, raw scores are scaled so that the subject average and the average for all students in all subjects in a particular semester of study are the same. These scores and grades are shown in columns 3 and 4, figure 7. In this program scaling does not equalise the standard deviations for the subject and for the year. Thus, there could be criticism that the scaling process does not go far enough. However, it does serve as a guide as is intended.

It should be stressed that this automatic scaling is done so

SUMMARY OF RESULTS.

| | | ACTUAL RES. | | RECOMMENDATION | ADJUSTED RES. | | SUBJECT | A.SCORE. |
|---|---|---|---|---|---|---|---|---|
| | | RES. | GRADE | | RES. | GRADE. | | |
| 13 | 11711730268A | 90.9% | ( 9 ) | ---------- | 70.7% | ( 7 ) | E.C.190 | 54.6% |
| | | 77.0% | ( 7 ) | ---------- | 66.0% | ( 6 ) | G.DRG130 | 54.6% |
| | | 76.0% | ( 7 ) | ---------- | 89.3% | ( 8 ) | MATH172 | 54.6% |
| | | 85.5% | ( 8 ) | ---------- | 84.1% | ( 8 ) | MDES130 | 54.6% |
| | | 75.0% | ( 7 ) | ---------- | 65.8% | ( 6 ) | MDRG130 | 54.6% |
| | | 71.0% | ( 7 ) | ---------- | 79.6% | ( 7 ) | MET.150 | 54.6% |
| | | 66.0% | ( 6 ) | ---------- | 69.9% | ( 6 ) | PHYS170 | 54.6% |
| | | 69.2% | ( 6 ) | ---------- | 63.5% | ( 6 ) | PIEC130 | 54.6% |
| | | 63.0% | ( 6 ) | ---------- | 72.6% | ( 7 ) | STAT.130 | 54.6% |
| | | 66.7% | ( 6 ) | ---------- | 90.1% | ( 9 ) | THER132 | 54.6% |
| | | 64.0% | ( 6 ) | ---------- | 64.8% | ( 6 ) | ENGL150 | 54.6% |

AVERAGE SCORE = 73.1% CORRECTED AVERAGE SCORE = 74.2%

**********

LEAVING CERTIFICATE RESULTS.

| ENGLISH | MATH.II | MATH.III | CHEMISTRY | PHYSICS |
|---|---|---|---|---|
| 66 | 63 | 76 | 74 | 75 |

MATRICULATION RESULTS.

| MATH.II | MATH.III | CHEMISTRY | PHYSICS. |
|---|---|---|---|
| 69 | 77 | -0 | 60 |

| | | ACTUAL RES. | | RECOMMENDATION | ADJUSTED RES. | | SUBJECT | A.SCORE. |
|---|---|---|---|---|---|---|---|---|
| | | RES. | GRADE | | RES. | GRADE. | | |
| 14 | 11711730182F | 75.0% | ( 7 ) | ---------- | 75.6% | ( 7 ) | DYN.130 | 54.6% |
| | | 62.9% | ( 8 ) | ---------- | 64.5% | ( 6 ) | E.C.190 | 54.6% |
| | | 93.5% | ( 9 ) | ---------- | 80.2% | ( 8 ) | G,DRG130 | 54.6% |
| | | 59.0% | ( 5 ) | ---------- | 69.3% | ( 6 ) | MATH172 | 54.6% |
| | | 68.0% | ( 6 ) | ---------- | 66.9% | ( 6 ) | MDES130 | 54.6% |
| | | 66.3% | ( 6 ) | ---------- | 74.3% | ( 7 ) | MET.150 | 54.6% |
| | | 73.0% | ( 7 ) | ---------- | 77.3% | ( 7 ) | PHYS170 | 54.6% |
| | | 76.2% | ( 7 ) | ---------- | 70.0% | ( 6 ) | PIEC130 | 54.6% |
| | | 83.0% | ( 8 ) | ---------- | 95.7% | ( 9 ) | STAT.130 | 54.6% |
| | | 66.7% | ( 6 ) | ---------- | 90.1% | ( 9 ) | THER132 | 54.6% |
| | | 58.0% | ( 5 ) | ---------- | 58.8% | ( 5 ) | ENGL150 | 54.6% |

AVERAGE SCORE = 72.9% CORRECTED AVERAGE SCORE = 74.8%

**********

| ENGLISH | MATH.II | MATH.III | CHEMISTRY | PHYSICS | MATH.II | MATH.III | CHEMISTRY | PHYSICS. |
|---|---|---|---|---|---|---|---|---|
| 58 | 70 | 72 | 69 | 76 | 68 | 65 | 66 | 70 |

FIG 7

that it is possible to assess overall performance meaningfully and to help in the development of the students future program. The Board of Examiners or the individual examiners are under no compulsion to use the scaled scores. They can elect to leave the score unchanged if deliberation suggests this is the wisest course. They can use the scaled score, or they can adopt another score as determined by the circumstances.

Figure 7 indicates the type of information provided to assist examiners in deciding on final scores. Consider case 13 in figure 7. The first column gives the raw score obtained by the student in each exam in the semester together with his average. The final column gives the semester average for all subjects and students in the identified semester subject group. The third column gives the scaled score and colum 4 the scaled grade. Thus for subject E.C. 190

$$\text{Scaled score} = \text{subject raw score} \times \frac{\text{average score for prescribed subjects in semester of study}}{\text{average for the subject}}$$

Finally the leaving and matriculation scores (alternate course entry examinations) are given.

to all Institute departments. Thus the first evaluation was restricted to the four departments of the School of Engineering and Surveying.

There has been a marked decline, to nil in 1974, in the number of disputed cases referred to arbitration. Again, departments which had opposed scaling in any form now accept that some form of scaling is required and this is done as a matter of course.

Originally the examining process was largely one of transferring raw examination scores in the shortest possible time from the department to the examinations office. Now, the best is made of the limited time available after the examinations[5] and judgements are made regarding the students' performance using all possible supporting information which is now made available.

Late in 1974 a report of the investigation, prepared by the writer and Mr. Nowak, was presented to the Institute's Academic Board which determines educational policy. This Board, in evaluating the report, agreed that

---

[5] In 1974 second semester examinations, the examination period extended over seven days followed by a period of nine working days before all material had to be submitted by examiners. All computer printouts for Boards of Examiners were available three actual days later.

"Academic Board commends the report to the working party on degrees with distinction, and recommends the procedures contained therein to interested departments."

It is hoped the procedure will be widely implemented.

The following points need further reviewing as the scheme is developed.

i. Mention has been made of the fact that the scaling process does not correct for variations in score spread. There could be value in ascertaining the effect of this second form of scaling.

ii. There are cases on record where scores in a majority of subjects in a particular year of a course have been depressed and investigation has suggested that this was due to reasons other than student ability. In the present scaling system used this could lead to erroneous results as these depressed semester scores are used to obtain the scaling average. It is suggested that more meaningful results could be obtained were the subject averages to be determined considering scores over a number of years.

iii. Attempts are being made to reduce the time of processing results by making use of a document reader. It is proposed that a pilot scheme should be introduced in 1975. The programme in its present form uses data from punched cards. As an alternate data can be read directly using the document reader. A major advantage here would be the elimination of the intermediate punching process in the transfer of information. This will reduce possible errors, expedite the process, and reduce the skilled labour component of the process thus effecting economies.

## 7 ACKNOWLEDGEMENTS

The author wishes to thank Mr. Nowak for the considerable time and application which he contributed in preparing the programs to further develop the work initiated by the author.

Thanks should also be expressed to the organisers of the Annual Conference for the American Society for Engineering Education for making it possible for this development in Australian engineering education to be presented to an audience in the United States of America thus helping to some extent in the sharing of ideas.

## REFERENCES

1. Nash, A.H. - Student Progress, Attitudes and Performance in Electrical Engineering Courses at the Western Australian Institute of Technology. 1969, Perth, Western Australian Institute of Technology, 157 p,